A Report from the University of Vermont Transportation Research Center

Development of GIS Tools to Optimize Identification of Road Segments Prone to Flood Damage

Final Report

TRC Report 15-005
September 2015

Ernie Buford, Jim Sullivan, Jarlath O'Neil-Dunne, and Austin Troy

Development of GIS Tools to Optimize Identification of Road Segments Prone to Flood Damage

Final Report

September 10, 2015

Prepared by:

Ernie Buford
Jim Sullivan
Jarlath O'Neil-Dunne
Austin Troy

Transportation Research Center
Farrell Hall
210 Colchester Avenue
Burlington, VT 05405

Phone: (802) 656-1312
Website: www.uvm.edu/trc

Acknowledgements

The authors would like to acknowledge VTrans for providing funding for this work.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the UVM Transportation Research Center. This report does not constitute a standard, specification, or regulation.

Table of Contents

1 Introduction

Highway drainage systems are designed to collect, convey, and discharge stormwater to prevent adverse impacts to the roadway or offsite. Most of these systems are engineered with a hydrograph in mind, but the magnitude of the structures statewide in Vermont prevents the Agency of Transportation (VTrans) from conducting a detailed analysis of each one. There are over 90,000 culverts tracked in the Vermont On-Line Bridge and Culvert Inventory Tool. Even when the structures are hydraulically engineered, drainage capacities can be exceeded during extreme precipitation events, resulting in road damage.

Roadway drainage infrastructure is designed in part based on the computation of runoff parameters and quantities (FHWA 2009). Therefore, it is not surprising that peak runoff flow volume has been found to be a key determinant in predicting flood damage to roads, more so than basin morphometric characteristics alone (Dawod et al. 2012). A number of methods exist for calculating runoff parameters and quantities. Among the most common approaches for larger watersheds (e.g. >200 acres), particularly where land cover is heterogeneous, is the SCS/NRCS Tabular Hydrograph method, which estimates partial composite flood hydrographs at any defined point in a watershed based on precipitation, drainage area, time of travel of water, land cover and other factors (SCS 1986). A few GIS tools have been developed in the past to evaluate the vulnerability to flood damage of roads specifically (e.g. Dawod et al. 2012) and of infrastructure in general, including roads (Karmakar et al. 2010, Zerger and Wealands, 2004). However, no such GIS tool is known to exist that is specifically designed for New England hydrologic environments, with their unique combination of hydrology, terrain, climate, land cover and geomorphology. The purpose of this project was to create an ArcToolbox interface for estimating peak runoff flow volume at culvert locations so that existing or planned infrastructure could be quickly assessed for flood-damage probability. A GIS-based tool was developed to compute runoff parameters at drainage infrastructure locations based on characteristics of input watersheds delineated at those locations.

This report presents the results of a geospatially-weighted regression analysis tool for estimating peak flow at culvert locations in Vermont, and describes the development of a GIS-based tool for quickly finding the parameters needed for the regression equation. Section 2 provides an overview of the data development and preprocessing steps conducted for the project, including watershed delineation and derivation of watershed characteristics used as inputs to perform the regression analysis. In Section 3, a detailed narrative description of the regressions analysis and the development of the GIS-based tool. The report concludes with a discussion of results and recommendations for applying and expanding the peak flow estimation tool.

2 Data Generation and Preparation

Datasets for the analysis were generated by delineating watersheds for a set of culvert locations with known peak flows and then overlaying the resulting polygons with geospatial data layers representing variables of interest for the regression. Terrain pre-processing, watershed delineation and derivation of watershed characteristics were accomplished using ArcHydro Tools and other geoprocessing tools in the ArcGIS Desktop (Versions 10.1 & 10.2) platform, as described below.

2.1 Culvert Locations and Peak Flows

VTrans conducts hydraulic analyses for culverts in Vermont that are being considered for repair or replacement. These analyses typically include peak flow determinations for several recurrence intervals and occasionally contain precise locational information in the form of measured latitude/longitude coordinates. Culvert locations and peak flows used in this study originated from the hydraulic reports that are completed for these analyses.

In the last decade, approximately 3,000 hydraulic analyses have been conducted, but only a small subset of those have specific coordinates for their precise location for geocoding. Reports for culverts in 221 Vermont towns were acquired and inspected to identify those that included both peak flow and location data. Latitude, longitude, peak flows, and descriptions were manually transcribed from nearly 300 hydraulic reports for use in this study. This data was used to create a shapefile of culvert locations with peak flow data attributes.

Point locations of culverts were further inspected in overlays with basemap layers (notably including stream lines) in a quality control process that ensured accurate alignment of culvert locations and hydrologic features that would determine contributing drainage areas. This step ensured that culvert points aligned with the water flow pattern represented in the DEM and stream channel datasets, so that subsequent hydrologic modeling would result in realistic watershed polygons.

The culvert points were augmented with additional attributes extracted from the shapefile of bridge and culvert inventory points obtained from the Vermont Center for Geographic Information. The culverts from this other shapefile that could be matched to the location descriptions contained in the hydraulic reports, helped add more culverts to the dataset where latitude and longitude were not available. Site photographs from VAOT, aerial imagery and other vector datasets guided the process of selecting the correct matching features. As a result of this process, the dataset of culvert locations with peak flows nearly doubled in size to a total of 581 points. These points are shown in Figure 1.

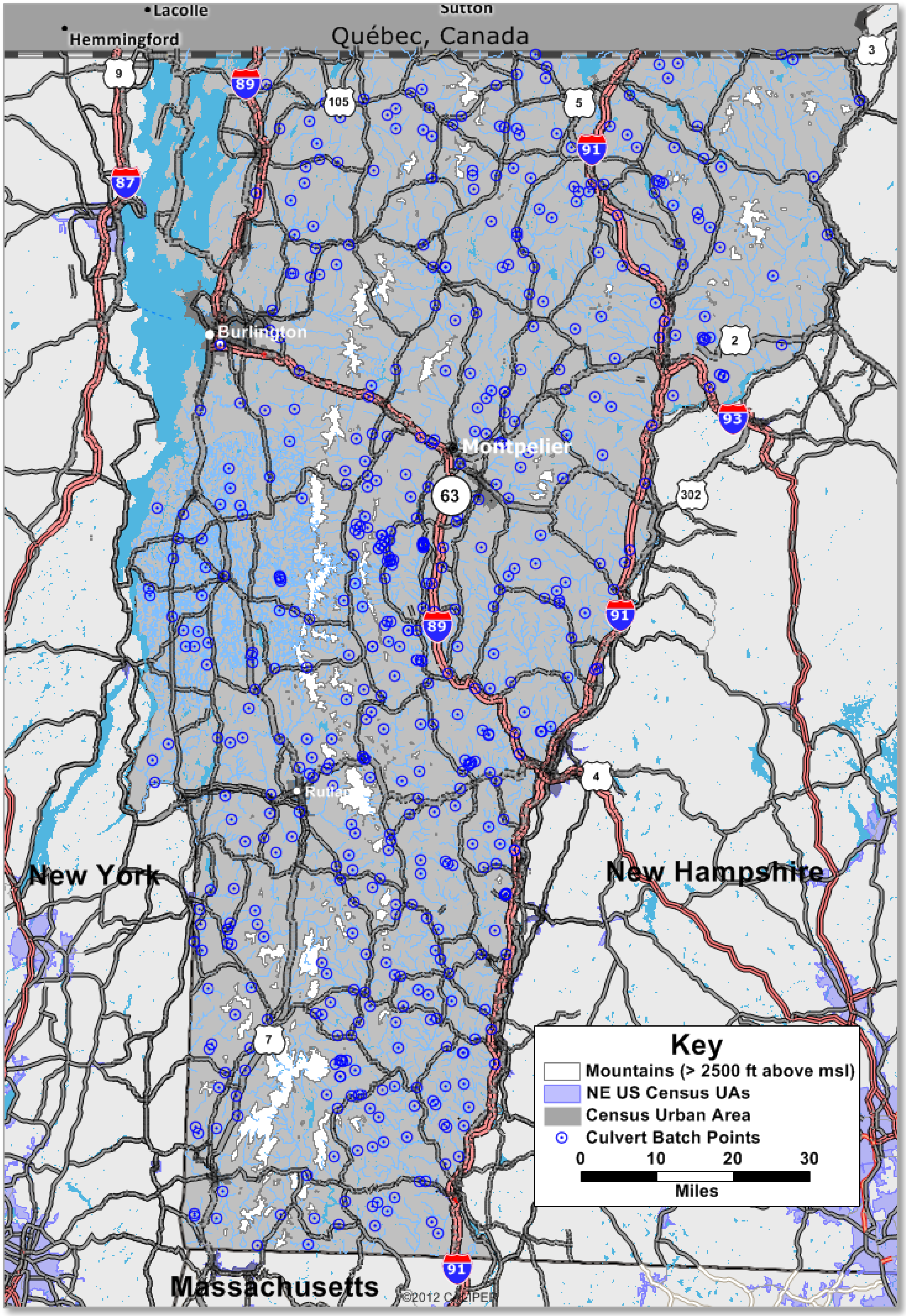

Figure 1 Culvert Batch Points Geo-Located for Use in This Study

2.2 Precipitation Data

Rainfall intensity levels and Bureau of Public Roads (BPR) rainfall factors were acquired from the VTrans Hydraulics Manual (VTrans, 1998). Contour maps of (1) 24-hour, 2-year maximum rainfall (2) 2-year, 30-minute rainfall intensity, and (3) BPR rainfall factors were scanned, georeferenced and digitized in ArcGIS to create vector datasets for each variable with the contour lines shown as features (Figure 2).

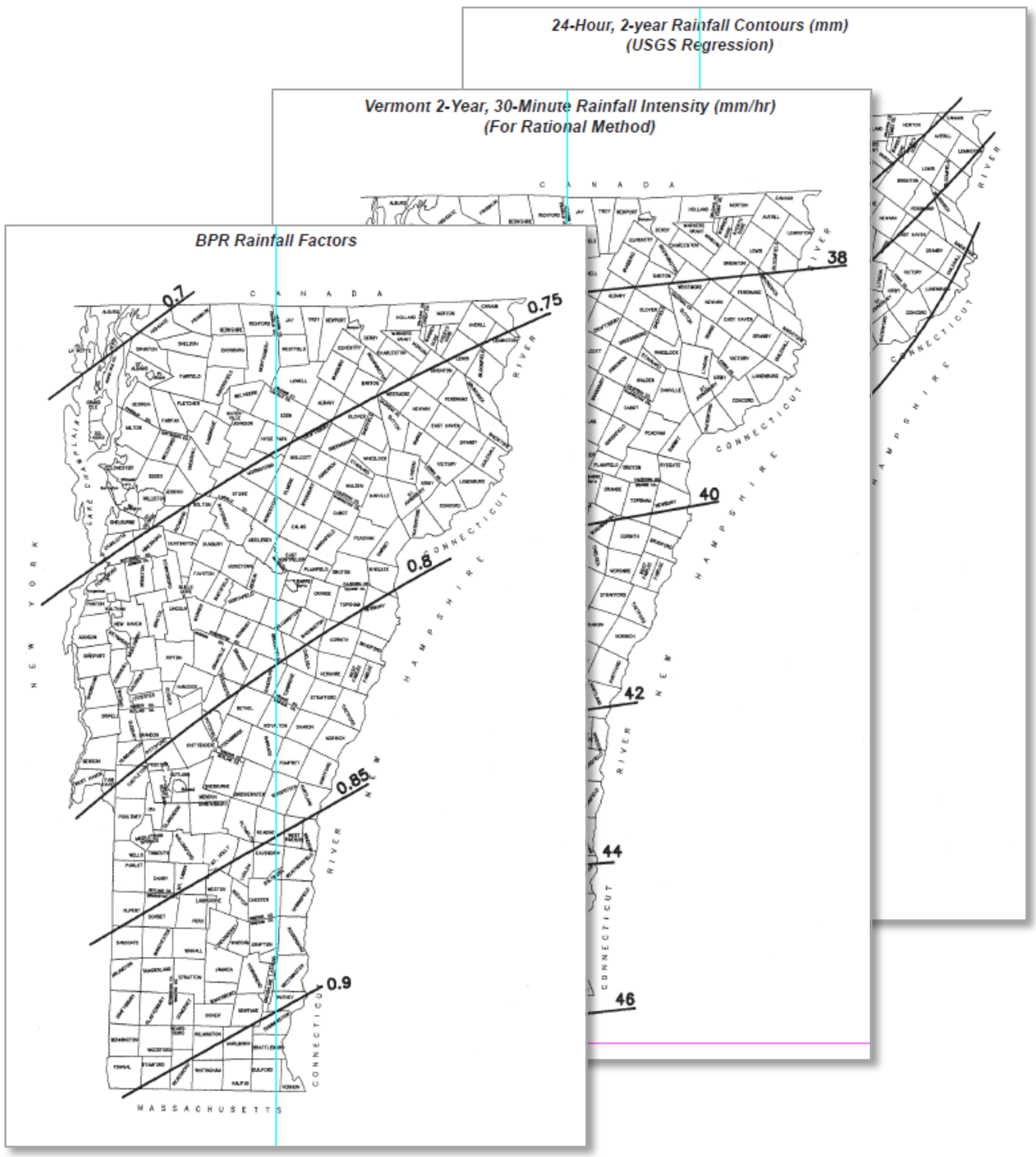

Figure 2 Contour Maps Scanned and Georeferenced for Use In this Study

Spatial interpolation was used to create a raster surface from each of these 3 vector datasets.

2.3 Land Cover Data

Land cover data for the regression analysis was taken from the 2006 National Land Cover Dataset (NLCD) (Fry et al., 2011) and the 2006 Coastal Change Analysis Program (C-CAP) (NOAA, 2006) rasters. The NLCD impervious surface raster was used without adjustment. Other land cover types used were wetlands (all "wetland"-related classes) and non-cultivated vegetation (pasture/hay, grassland/herbaceous, scrub/shrub and all vegetated, palustrine/estuarine classes). For these two land cover types, the C-CAP raster was reclassified to a binary raster with pixels coded as 1 (subject cover type present) or 0 (other cover type present).

2.4 Stream Flowlines

Stream flowlines were taken from the National Hydrography Dataset (NHD) Version 2.1. NHD geodatabases for Vermont watersheds were acquired from USGS (2013). Each geodatabase covers one Hydrologic Unit Code (HUC) level 4 drainage area. The HUC 4 areas covering Vermont are the Merrimack, Connecticut, Upper Hudson and St. Lawrence river basins. NHD flowlines are useful in representing stream centerlines for hydrologic modeling procedures and hierarchical watershed boundaries that have utility in partitioning an area into smaller hydrologic units for data extraction and analysis. Flowlines were extracted for all watersheds covering Vermont.

2.5 Digital Elevations

A digital elevation model (DEM) for Vermont was acquired from the Vermont Center for Geographic Information (VCGI). This DEM (ElevationDEM_DEM10M), with a 10-meter pixel size, represented the best available, statewide DEM data for watershed delineation. However, since watersheds for culverts near state boundaries can extend into neighboring states or into Quebec, Shuttle RADAR Topography Mission (SRTM) (EROS, 2009) elevation data were used to augment the DEM to accommodate delineation of culvert watersheds that cross the state boundary. Z units of the augmented DEM were converted to centimeters as recommended for use in Arc Hydro.

2.6 Terrain and Watershed Preprocessing

A terrain preprocessing step was performed to create a hydrologically-correct DEM in order to prepare it for the hydrologic analysis. Terrain preprocessing was performed separately for 9 major river basins, excluding several small areas draining directly into Lake Champlain rather than into one of the major rivers. These basins were constructed from NHD watershed boundaries. The workflow chosen for delineation and analysis of Vermont culvert watersheds was a modified version of one developed by the makers of the ArcGIS software (ESRI, 2013) This process was informed by additional documentation of terrain preprocessing methods (Djokic, 2008). A combination of Arc Hydro geoprocessing models and custom ArcGIS models automated the majority of these terrain preprocessing steps for each river basin. The resulting workflow included the following steps:

1 Extract NHD flowlines, DEM and culvert locations for a basin
2 Manually edit NHD flowlines to delete non-outlet line features that cross the basin boundary
3 Set up a workspace and map document for the basin, set target (output) locations and assign Hydro ID numbers to NHD flowlines
4 Create drainage line structures
5 DEM reconditioning using parameters:
    a. Buffer: 5 cells
    b. Smooth drop/raise distance: 140 cm
    c. Sharp drop/raise distance: 1000 cm
6 Fill sinks
7 Derive flow direction
8 Adjust flow direction in streams
9 Derive flow accumulation
10 Delineate catchment grid
11 Catchment polygon processing
12 Adjoint catchment processing

After completing the terrain preprocessing step, watershed delineation was conducted for the regression analysis using the ArcHydro "Batch Watershed Delineation" tool. This processing step used the adjusted flow-direction raster and rasterized flowlines as inputs to delineate watershed polygons for the 581 culvert point locations. The tool "snapped" culvert point locations two pixels (20-meter) or less away to coincide with a flowline (points farther away from a flowline were not snapped).

Polygons resulting from the "Batch Watershed Delineation" tool were manually inspected in overlays with flowlines and other data layers to check for errors. Culvert point locations were manually adjusted in cases where delineated watersheds were obviously unrealistic, and "Batch Watershed Delineation" was repeated using the adjusted point locations. For some basins, this process was repeated in an iterative process (as many as 5 times for one basin) until satisfactory results were achieved.

Additional ArcHydro tools were applied to calculate additional hydrologic and topographic characteristics of the delineated watersheds for the regression. These characteristics included flow path length, average slope, total vertical drop, and slope & elevation at locations 10% and 85% along the flow path.

Another custom ArcGIS geoprocessing tool was developed to overlay the 581 culvert watersheds with datasets representing the set of independent variables. The tool developed for this task used the "Zonal Statistics as Table" tool to summarize each variable in each watershed polygon and then compile all variables in an output table suitable for regression analysis:

1. Total watershed area (in square meters)
2. Length of longest flow path (in miles)

3. Average flow-path slope (in feet per mile)
4. Average slope between 10% and 85% of the flow path (feet per mile)
5. Upstream elevation (in meters above mean sea level)
6. Downstream elevation (in meters above msl)
7. Elevation at 10% along flow path from outlet (in meters above msl)
8. Elevation at 85% along flow path from outlet (in meters above msl)
9. Elevation range of watershed (in centimeters)
10. Average elevation of watershed (in centimeters above msl)
11. Average percent of watershed area impervious surface
12. Average percent of watershed area classified as vegetated (non-cultivated)
13. Average percent of watershed area classified as wetland
14. BPR Rainfall Factors
15. 24-hour, 2-year rainfall (in mm)
16. 30-minute, 2-year rainfall intensity (in mm per hour)

3 Methodology and Results

3.1 Regression

Ordinary least-squares (OLS) regression was used to develop a regression equation for predicting peak flows at a given culvert using characteristics of the culvert's watershed. Backwards, stepwise regression was performed on the culvert watershed characteristics dataset using peak flows at 100-year recurrence interval as the dependent variable and the list of variables described in Section 2.6 as the independent variables.

Several steps were conducted in an effort to develop the best-fit to a regression equation. Final sample size of the input culvert dataset was reduced by approximately 100 culverts in an effort to minimize spatial autocorrelation in the data. Log-transformation of peak flow and watershed area greatly improved the predictive power of the regression model, which resulted in an adjusted R-squared of 0.69 and the following equation:

$$\ln Q = 1.332 - 0.001S + 0.00001658Z + 0.022R + 0.188L + 0.147 \ln A - 0.018W$$

where:

Q = peak flow, 100-year recurrence interval (cubic feet per second)

S = channel slope (feet per mile)

Z = elevation range (cm)

R = maximum 24-hour rainfall, 2-year recurrence interval (mm)

L = flow path length (mi)

A = watershed area (square meters)

W = percent wetland land cover

The 9 other watershed attributes described in Section 2.6 were not found to contribute significantly to the regression. These other attributes were either insufficiently related to the peak flows, or they were correlated with one of the 7 attributes used in the equation to the extent that they were not needed in the equation.

3.2 GIS-Based Tool Development

An ArcGIS Desktop geoprocessing tool was created by developing a Python script to compute peak flow and other characteristics for a given set of input watershed polygons generated for culvert locations of interest. This tool expanded on a combined iteration and batch-processing approach initiated for the custom geoprocessing tool described above (Section 2.6). The iteration approach was only necessary for input watersheds that overlap each other. The combined approach (iteration and batch-processing) allowed efficient processing of non-overlapping watershed polygons while maintaining ability to process overlapping watersheds that would not be handled properly if watersheds were done in batch.

The tool was designed to take as an input a feature class of pre-generated watershed polygons for one or more culvert locations. Other required inputs were raster datasets for deriving variables needed to compute peak flow: (1) the DEM used to generate watersheds, (2) a binary representation of wetland land cover (as described in Section 2.3), (3) maximum 24-hour rainfall, 2-year recurrence interval (Section 2.1), and (4) flow path length in meters. Script development accommodated the possibility that input watersheds may have been generated by ArcHydro. Raster processing steps required the use of tools from the Spatial Analyst extension to ArcGIS Desktop.

At the most recent stage of development, the tool derived watershed characteristics needed for the regression equation and computed peak flow along with a bank full width estimate desired by state transportation engineers. Outputs were designed to be written to a geodatabase table for joining to input watershed polygons or source culvert locations having the necessary matching ID.

4 Conclusions and Recommendations

Geospatially-weighted regression delivers improvements over statistical methods that do not account for spatial non-stationarity of relationships between variables. This project successfully derived a new regression equation that provides predictive power for peak flow estimation at all bridge and culvert locations in Vermont, accounting for the state's unique topographic characteristics. This regression model relied heavily on accurate delineation of watersheds for bridge and culvert locations. Delineating watersheds for bridge and culvert locations was automated using a GIS tool with reasonable results for a majority of the state's structures, particularly those located on major rivers and streams represented in the drainage network. Many other structures, however, required manual adjustment of pour-point locations to produce realistic watershed boundaries and realistic regression coefficients.

The peak flow estimation tool developed in this project is flexible enough to be coupled with existing watershed delineation techniques in a larger workflow for dynamically generating watersheds, deriving watershed characteristics and computing estimates of flow conditions at all of Vermont's 90,000+ culverts and bridges. The number of culverts addressed simultaneously depends on processing constraints (hardware and software limitations) and the level of effort necessary for manual quality control adjustments. This estimation was not feasible at this point in time due to the limited amount of calibration data that was available when this project was conducted.

Accuracy of the terrain model (DEM) and potential, scale-dependent alignment issues with respect to culvert locations is a persistent challenge for accurate, automated watershed delineation. Snapping culvert locations to be aligned with the nearest stream channel can improve automatic watershed delineation results significantly. However, this solution would not improve watershed delineation for points not within the snapping tolerance of an existing river or stream. For example, drainage channels along roadsides present significant challenges to this procedure. Snapping can also produce erroneous results for locations away from flowlines, where the DEM may poorly represent actual terrain, and for culverts that happen to be very near but not coincident with flowlines. Future research should focus on resolving this problem, since many of the flow-related damage points from Hurricane Irene occurred within drainage channels and ditches that do not normally carry water.

Given the modest explanatory power of the peak-flow regression model developed in this project, further validation of predicted peak flows, where data can be acquired, would be helpful in evaluating model performance as an aid to evaluation or design of culverts. However, the fit of the model developed in this project (r-squared of 0.69) compares well with the fit of other, non-spatial models focused on Vermont (Olson, 2002).

Once the regression model that was developed in this project is improved, additional validation can be conducted at road segments designated to have a high probability of flooding against road damage data for Tropical Storm Irene. This validation will quantify the extent to which the designation of risk assigned by a peak-flow model works to predict the spatial distribution of road damage from a major precipitation event like Hurricane Irene. It may also be possible at that time to fine-tune the regression model to maximize its predictive power in forecasting the likelihood of road damage throughout the state in an automated fashion.

5 References

Dawod, G., Mirza, M. and Al-Ghamdi, K. 2012. GIS-based estimation of flood hazard impacts on road network in Makkah city, Saudi Arabia. Environmental Earth Sciences. On-line preprint.

Federal Highway Administration (FHWA). 2009. Urban Drainage Design Manual. Hydraulic Engineering Circular No. 22, Third Edition. Washington, DC.

Karmakar, S., Simonovic, S., Peck, A. and Black, J. 2010. An information system for risk-vulnerability assessment to flood. Journal of Geographic Information System, 2(2): 129-146.

Soil Conservation Service (SCS), 1986. Urban Hydrology for Small Watersheds, Technical Release No. 55, U.S. Department of Agriculture

Zerger A,and Wealands S. 2004. Beyond modelling: linking models with GIS for flood risk management. Natural Hazards, 33:191–208.

Djokic, D. 2008. Comprehensive Terrain Preprocessing Using Arc Hydro Tools. Accessed September 2013 at https://mft.esri.com.

(EROS) U.S. Geological Survey, Earth Resources Observation and Science Center. 2000. Shuttle RADAR Topography Mission Data. Accessed August 2010 at http://eros.usgs.gov.

ESRI, 2013. Overview of Arc Hydro Terrain Preprocessing Workflows. Accessed September 2013 at https://mft.esri.com.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States, PE&RS, Vol. 77(9):858-864.

National Oceanic and Atmospheric Administration (NOAA), Office for Coastal Management. "2006 C-CAP Regional Land Cover." Coastal Change Analysis Program (C-CAP) Regional Land Cover. Charleston, SC: NOA Office for Coastal Management. Accessed December 2008 at http://www.coast.noaa.gov/ccapftp.

Olson, S. A. 2002. Flow Frequency Characteristics of Vermont Streams. Accessed September 2013 at http://pubs.usgs.gov/wri/wrir02-4238/.

USGS, 2013. National Hydrography Dataset. Accessed September 2013 at http://nhd.usgs.gov.

VTrans, 1998. Hydraulics Manual. Prepared by the Vermont Agency of Transportation.